基于聚类 PSO-LSSVM 模型的 PAD 维度预测 *

胡艳香，孙 颖†，张雪英，段淑斐

(太原理工大学 信息与计算机学院，太原 030024)

摘 要：针对 PAD(pleasure：愉悦度、arousal：激活度、dominance：优势度)预测精度问题，提出将最小二乘支持向量机(least squares support vector machine，LSSVM)经粒子群优化(particle swarm optimization，PSO)算法优化再与情感聚类分析结合的聚类 PSO-LSSVM 模型。对 TYUT2.0 和柏林语音库的三种情感语音提取情感特征，基于特征与标注的 P、A、D 对三种单一情感分别建立各类情感维度 PSO-LSSVM 模型以及对三种情感建立混合情感维度 PSO-LSSVM 模型；然后利用混合情感维度 PSO-LSSVM 模型预测 P、A、D，并计算其与基本情感 PAD 的距离；最后将距离大于阈值的情感聚类为混合情感，将距离小于阈值的情感聚类为与其距离最近的情感，并利用对应情感的回归模型预测其 P、A、D。研究显示，该模型对 P、A、D 的预测误差较 LSSVM 和 PSO-LSSVM 模型更小，且预测值与标注值的相关性更强，说明聚类 PSO-LSSVM 模型对 P、A、D 的预测更加可靠、准确。

关键词：情感维度 PAD；最小二乘支持向量机；粒子群优化算法；情感聚类分析

中图分类号：TN912.3 doi: 10.19734/j.issn.1001-3695.2018.10.0735

Forecast of pad dimensions using clustering pso-lssvm model

Hu Yanxiang, Sun Ying+, Zhang Xueying, Duan Shufei

(*College of Information & Computer*, *Taiyuan University of Technology*, *Taiyuan* 030024, *China*)

Abstract: In view of the imprecision problem for PAD(Pleasure、Arousal、Dominance) prediction, this paper proposed clustering PSO-LSSVM model combineing Least Squares Support Vector Machine(LSSVM) optimized by Particle Swarm Optimization(PSO) and affective clustering analysis. Firstly, selecting three emotion speeches of TYUT2.0 emotional speech database and Berlin voice library, and extracting emotion features. Establishing Single emotional dimension PSO-LSSVM models for three single emotion and the mixed emotion dimension PSO-LSSVM model for three emotions based on emotion features and P, A and D values. The mothod used mixed emotion dimension PSO-LSSVM model to predict the P, A and D values of the test set, and calculated the distance between the predictive PAD and the PAD of the basic emotion. Finally clustering the emotion whose distance is greater than the threshold into mixed emotion, and clustering the emotion whose distance is less than the threshold into the nearest emotions, then using the corresponding emotional dimension regression model to predict its P, A and D. The research showed that the predictive error of clustering PSO-LSSVM regression model to P, A and D was smaller than that of LSSVM and PSO-LSSVM model, and the correlation between the predicted value and the tagged value was stronger. So the clustering PSO-LSSVM regression model is more reliable and accurate in predicting P, A and D values.

Key words: emotional dimensions pad; least squares support vector machine; particle swarm optimization algorithm; affective clustering analysis

0 引言

情感计算对于成熟的人工智能发展是不可或缺的，情感维度从情感的心理学属性对情感进行描述，在情感计算[1]方面具有重要意义。随着人机交互的智能化发展，离散情感已经难以满足生活化的自然情感识别，因此研究蕴涵情感信息的连续情感维度就显得尤为重要，鉴于情感维度难以实时监测的特点，决定了要通过建立数学模型来进行维度预测，而对数据进行模拟与拟合是进行有效预测的主要手段[2]。

人工智能预测方法主要包括神经网络和支持向量机。但神经网络所需数据样本大，在小样本的情感语音中无法保证预测精度。支持向量机(support vector machines, SVM)[3]可有效解决小样本、非线性等复杂系统问题，但需要求解二次规划方程，计算复杂性较大。最小二乘支持向量机(least squares support vector machine,LSSVM)[4]通过等式约束对 SVM 进行改进，同时参数选择优化方法即粒子群优化(particle swarm optimization, PSO)算法[5]的引入使其建模过程中的参数选择更为客观。PSO-LSSVM 算法目前主要应用于太阳能温室温度预测[6]、节流液速预测[7]等方面，尚未见有资料将其应用于情感维度预测。但是不同情感的情感特征之间错综复杂的相互作用容易影响 PSO-LSSVM 回归模型对情感维度的预测精度，因此在该回归模型中引入情感聚类分析以降低不同情感情感特征之间的影响是必要的。本文提出根据 PSO-LSSVM 模型的初次预测结果与基本情感中心 PAD 的距

收稿日期：2018-10-18；**修回日期：**2018-11-22 **基金项目：**国家自然科学基金资助项目（61371193）；山西省青年基金项目（2013021016-2）；山西省研究生教育创新项目（2018SY021）；山西省应用研究青年基金资助项目（201601D202045）

作者简介：胡艳香（1993-），女，山西五台人，硕士研究生，主要研究方向为情感计算；孙颖（1981-），女，（通信作者），山西太原人，讲师，硕导，博士，主要研究方向为情感语音识别、情感计算（tyutsy@163.com）；张雪英（1964-），女，山西太原人，教授，博导，博士，主要研究方向为语音信号处理、情感识别；段淑斐（1983-），女，山西清徐人，讲师，硕导，博士，主要研究方向为声音信号处理及动物声音识别.

离进行聚类处理，再采用 PSO-LSSVM 回归模型进行二次维度预测。

本文以 PAD 三维情感模型[8]为基础，首先提取情感语音的情感特征；然后分别采用三组回归模型：LSSVM 回归模型、PSO-LSSVM 回归模型、聚类 PSO-LSSVM 回归模型建立语音的情感特征(韵律特征、MFCC 特征和非线性特征)与人工标注 P、A、D 值的映射关系，实现对情感语句的 P、A、D 值预测；最后对三类回归模型的预测结果进行比较分析，实验结果表明聚类 PSO-LSSVM 回归模型对 P、A、D 值的预测精度更高。

1 PAD 三维情感模型

日常生活中人类的情感比较微妙和复杂，如悲喜交加、喜极而泣等情感不完全属于某一基本情感类别，故提出情感的连续空间论[9]来解决该问题。该理论认为人类的情感由空间的几个维度组成，该空间几乎可以涵盖人类所有的情感，不同情感可以实现连续、平稳地转变。连续情感模型中较为典型的是 PAD 三维情感模型，该模型由 UCLA 大学的 MEHRABIAN 开发[8]，采用语义差异评价方法将情感分为三个维度，分别为 P 代表愉悦度（pleasure-displeasure，表示个体情感状态的正负特性）、A 代表激活度（arousal-nonarousal，表示个体的神经生理激活程度）和 D 代表优势度（dominance-submissiveness，表示个体对情境和他人的控制状态）。PAD 三维情感模型从 P、A、D 三个角度对情感进行描述，便于量化日常生活中连续、多样的情感。PAD 三维情感模型如图 1 所示。

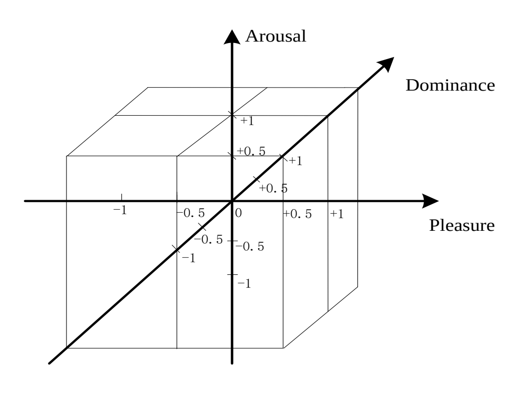


图 1 PAD 三维情感模型

Fig. 1 PAD dimensions model

2 理论基础

2.1 最小二乘支持向量机理论

LSSVM 算法[4]的基础是支持向量机(support vector machine,SVM)。SVM 将输入的数据通过非线性映射方式映射到高维特征空间，将问题转换为一个由不等式约束的二次规划问题，但其计算复杂度较大，故将最小二乘线性理论引入到 SVM 中进行改进，使得传统的二次规划问题转换为求解线性方程组问题，降低计算的复杂度。LSSVM 算法基本原理是：

设定样本集：

$$\{x_i, y_i\}, i=1,2,\cdots,n \begin{cases} \min \quad f(\mathrm{C},\sigma)=\sum_{i=1}^{n}(y_i-\hat{y})^2 \\ s.t. \quad C\in[\mathrm{C}_{\min,}\mathrm{C}_{\max}], \sigma\in[\sigma_{\min},\sigma_{\max}] \end{cases}$$

其中：$x_i \in R^d$ 为输入变量；$y_i \in R$ 为相应的输出；n 为样本集的大小；$\Phi(x)$ 将低维样本映射到高维特征空间 H ,构造最优决策函数 $y=\omega^T\Phi(x)+b$ (w 是权向量，b 是偏差)。对于外输入的样本 x，有 $|y-w^T\Phi(x)-b|\le e_i$，则 LSSVM 的优化问题为

$$\begin{aligned} &\min J(w,e)=\frac{1}{2}w^T w+\frac{1}{2}C\sum_{i=1}^{n}e_i^2 \\ &s.t. \quad y_i=w^T\Phi(x_i)+b+e_i,(i=1,2,\cdots,n) \end{aligned} \tag{1}$$

其中：C 为正则化参数；$e_i \in R$ 为误差变量。用 Lagrange 法求解以上优化问题，转换为以下线性问题：

$$\begin{bmatrix} 0 & \boldsymbol{I}^T \\ \boldsymbol{I} & \boldsymbol{K}+\boldsymbol{I}/C \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix}=\begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix} \tag{2}$$

其中：$\boldsymbol{I}=[1,1,\cdots,1]^T$ 为 n 个 1 组成的向量，为 n 阶单位矩阵；$\boldsymbol{y}=[y_1,y_2,\cdots,y_n]^T$；$\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\cdots,\alpha_n]^T$ 为 lagrange 乘子向量；$\boldsymbol{K}$ 为核函数矩阵，$K(x_i,x_j)=\Phi(x_i)^T\Phi(x_j), i,j=1,2,\cdots,n$。本文实验采用径向基(RBF)核函数，$\sigma$ 是核函数宽度，RBF 核函数为 $K(x,x_i)=\exp(-\frac{\|x-x_i\|^2}{2\sigma^2})$。

最后得到 LSSVM 模型：

$$y(x)=\sum_{i=1}^{n}\alpha_i \boldsymbol{K}(x,x_i)+b \tag{3}$$

在 LSSVM 模型中，x 表示模型输入，$y(x)$ 表示模型输出，在本文研究中，x 代表语音情感特征，$y(x)$ 表示语音情感维度值。根据 LSSVM 的回归原理可知，有两类参数的选择是需要特别关注的，即正则化参数 C 和核参数 σ。为减少对两类参数主观选择的盲目性，本文采用粒子群优化的方法来选取 LSSVM 的正则化参数 C 和核参数 σ。

2.2 粒子群优化算法

PSO 算法[5]是一种基于群体智能的优化计算方法。该算法在对飞鸟集群行为观察的基础上，利用群体中的个体对信息的共享，实现整个群体从无序到有序的演化运动过程，从而得到最优解。

PSO 算法中，$u_i=(u_{i1},u_{i2},\cdots,u_{in})$ 表示粒子 i 的位置向量，$p_i=(p_{i1},p_{i2},\cdots,p_{in})$ 表示第 i 个粒子经历过的最好位置，每个粒子的速度向量表示为 $v_i=(v_{i1},v_{i2},\cdots,v_{in})$，种群中所有粒子经历过的最好位置为 $p_g=(p_{g1},p_{g2},\cdots,p_{gn})$，粒子 i 的速度、位置更新公式如下：

$$\begin{cases} v_{in}^{k+1}=wv_{im}^{k}+c_1 r_1(p_{in}-u_{in}^{k})+c_2 r_2(p_{gn}-u_{in}^{k}) \\ u_{in}^{k+1}=u_{in}^{k}+v_{in}^{k+1} \end{cases} \tag{4}$$

其中：r_1 和 r_2 是 [0,1] 之间的随机数；w 是惯性权重，可以平衡算法的全局搜索能力以及局部搜索能力；c_1 和 c_2 分别是自我学习因子和社会学习因子。

2.3 基于 PSO 优化的 LSSVM 回归模型

定义目标函数：

$$\begin{cases} \min \quad f(\mathrm{C},\sigma)=\sum_{i=1}^{n}(y_i-\hat{y})^2 \\ s.t. \quad C\in[\mathrm{C}_{\min,}\mathrm{C}_{\max}], \sigma\in[\sigma_{\min},\sigma_{\max}] \end{cases} \tag{5}$$

其中：y_i 是第 $\boldsymbol{i}$ 个情感样本的 P、A、D 维度标注值；$\hat{y}_i$ 是模型对样本的 P、A、D 预测输出值，可由式(3)计算得到。PSO 算法优化 LSSVM 模型的目标是通过一系列迭代算法搜索一组参数 (C,σ)，使目标函数式(5)，即情感维度的主观标注值与模型对维度的客观预测值的误差达到最小。

2.4 情感聚类分析

在语音情感聚类中，传统的聚类方法是在特征空间上进行聚类[10]，这种聚类结果产生的依据是语音信号，然而这样的聚类结果未必与情感相关。针对语音情感聚类问题，本文通过描述连续情感的情感维度进行聚类，即通过语音的初步 PAD 预测值与基本情感中心 PAD 值的距离进行情感聚类。其流程如图 2 所示。

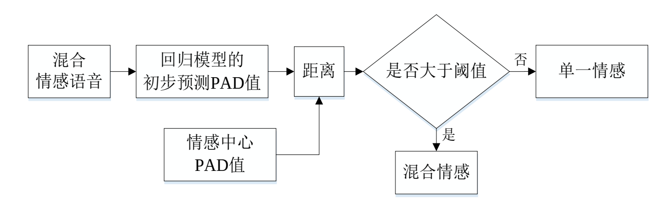


图 2　情感聚类分析流程

Fig. 2　Flow chart of emotional clustering analysis

尽管聚类算法很多，如 k-均值聚类、模糊聚类等，但都没有可以直接确定聚类阈值的方法。为了直观地确定聚类阈值，阈值的确定依据情感语音的 PAD 初步预测值与基本情感的欧氏距离[11]进行取值。欧氏距离计算表达式为

$$\mathrm{dist}(Y,Z)=\sqrt{(\hat{y}_P-z_P)^2+(\hat{y}_A-z_A)^2+(\hat{y}_D-z_D)^2} \tag{6}$$

其中：$\hat{y}_P$、$\hat{y}_A$、$\hat{y}_D$分别为情感语音的 P、A、D 预测值；z_P、z_A、z_D分别为基本情感的 P、A、D 值。

情感聚类分析将通过首次对语音的 PAD 预测值与基本情感中心的 PAD 值进行距离分析，根据距离与阈值的关系将语音区分为特定的单一情感和混合情感，在二次预测时将语音采用对应情感的 PAD 回归模型进行预测。

2.5　聚类 PSO-LSSVM 回归模型

在情感语音样本中，情感本身之间的关联性使不同情感语音的情感特征之间也存在相关性，在预测 P、A、D 的过程中会影响回归模型的预测性能，因此提出将 PSO-LSSVM 与情感聚类分析相结合的回归方案。图 3 所示为本文提出的聚类 PSO-LSSVM 回归模型流程。

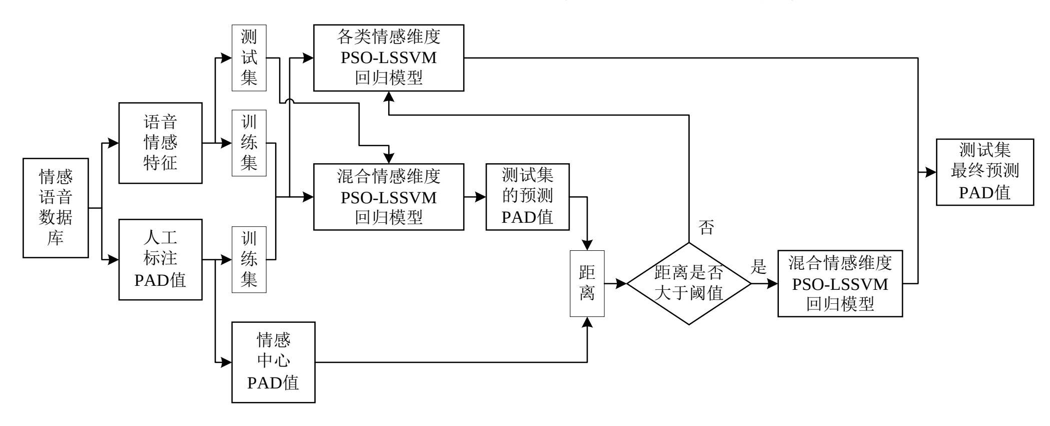


图 3　聚类 PSO-LSSVM 回归模型流程

Fig. 3　Flow chart of clustering PSO-LSSVM regression model

具体如下：首先对情感语音提取情感特征，基于训练集的情感特征和标注 P、A、D 值建立两类回归模型，一类称为各类情感维度 PSO-LSSVM 回归模型，此模型由单种情感语音的情感特征与人工标注 P、A、D 值对 PSO-LSSVM 训练得到，一类称为混合情感维度 PSO-LSSVM 回归模型，此模型由多种情感语音的情感特征与人工标注 P、A、D 值对 PSO-LSSVM 训练得到；然后将测试集的语音情感特征作为混合情感维度 PSO-LSSVM 模型的输入变量预测得到 PAD 值，并计算其与基本情感的中心 PAD 之间的距离，基本情感的中心 PAD 由对各类情感的 PAD 标注结果进行模糊 *C* 均值聚类得到，将距离大于阈值的情感聚类为混合情感，将距离小于阈值的情感聚类为与其距离最小的情感；最后将聚类为混合情感的语音情感特征作为混合情感维度 PSO-LSSVM 回归模型的输入变量，该模型与由训练集训练得到的混合情感维度 PSO-LSSVM 回归模型是同一个回归模型，而将聚类为某种特定情感的语音的情感特征作为由训练集训练得到的各类情感维度 PSO-LSSVM 模型中对应情感的模型输入变量预测其 PAD 值，该过程使得测试集的语音通过各自对应回归模型预测其 PAD，降低了不同情感特征间的相关性对预测效果的影响。该回归模型由 PSO 优化的 LSSVM 与聚类分析相结合，不仅避免了回归过程参数选择的主观盲目性，而且降低了输入变量之间的相关性，可以对情感维度 P、A、D 实现更精确的预测。

3　实验与分析

3.1　实验流程

为了验证聚类 PSO-LSSVM 回归模型在情感维度预测的有效性，设计了三组回归模型进行对比，即将 LSSVM 回归模型(模型一)、PSO-LSSVM 回归模型(模型二)、聚类 PSO-LSSVM 回归模型(模型三)的实验结果进行对比。实验流程如下：

a)对 TYUT2.0 数据库情感语音提取情感特征；

b)基于提取的情感特征和人工标注的 P、A、D 值，利用三组回归模型(LSSVM，PSO-LSSVM，聚类 PSO-LSSVM)预测测试集的情感维度 P、A、D；

c)对不同回归模型对 P、A、D 的预测结果与人工标注结果进行分析，选择更加合理有效的回归模型。

3.2　实验数据

3.2.1 数据库

情感语音数据库是进行语音情感分析的重要前提。为了全面、客观地评价模型对 P、A、D 的预测能力，选用 TYUT2.0 中文情感语音数据库[12]和柏林德语情感语料库(EMO-DB)[13]中共有的情感类型：悲伤(52 句)、愤怒(57 句)和高兴(52 句)作为实验样本，其中 67%作为训练样本，33%作为测试样本。数据库的情感语音用于提取特征，并将情感特征和每句语音对应的 PAD 值用于训练回归模型；根据每句语音的 PAD 预测值与基本情感的 PAD 进行聚类分析。

TYUT2.0 情感语音公开数据库是通过截取广播剧的方式获得的一种摘引型情感数据库，并根据改进的 PAD 情感量表对该数据库的语音进行维度标注，得到每句语音对应的 P、A、D 值(可联系作者邮件获取)。

柏林情感语音库由 10 位演员对 7 种情感模拟得到。该数据库的语音录制要求演员通过回忆自身经历来完成情绪的表达，使得语音情感真实度高、使用度较广、代表性较强。

3.2.2 语音情感特征

为对语音情感进行较完整的表征，并实现对 P、A、D

值更精确地预测，选取的特征如表 1 所示。

表 1　语音情感特征

Table 1　Emotional speech characteristics

特征属性	特征名称
MFCC	MFCC 前 12 阶的偏度、峰度、均值、方差、中值
韵律特征	语速 平均过零率 能量及其 1 阶差分的最大值、最小值、均值 基频及其 1 阶差分的最大值、最小值、均值 第 1 共振峰及其 1 阶差分的最大值、最小值、均值、方差 第 2 共振峰及其 1 阶差分的最大值、最小值、均值、方差 第 3 共振峰及其 1 阶差分的最大值、最小值、均值、方差
非线性特征	Hurst 指数的最大值、最小值、均值、中值、方差 最小延迟时间的最大值、最小值、均值、中值、方差 关联维数的最大值、最小值、均值、中值、方差 Kolmogorov 熵的最大值、最小值、均值、中值、方差 最大 Lyapunov 指数的均值、中值、方差

如表 1 所示，本实验从两个角度选择特征，针对语音信号的短时平稳特性提取声学特征，即 MFCC 特征(60 维)和韵律特征(38 维)，以及针对语音的混沌特性提取非线性特征[14](23 维)，将它们融合后，得到 121 维的特征集。

3.3　聚类分析结果

情感聚类分析的目的是提高对语音的 PAD 预测精度。为了保证情感聚类距离阈值的取值不受单一数据库的影响，通过对 TYUT2.0 数据库和柏林数据库两个数据库的语音 PAD 预测值与基本情感的 PAD 进行聚类分析统计，选择使聚类效果较好的聚类距离。根据情感聚类距离不同，不同聚类阈值区间内包含的情感语音数目如表 2 所示。

表 2　不同聚类阈值区间内的语音数目

Table 2　Number of speech in different cluster thresholds

dist²	TYUT2.0 数据库	柏林语音库
[0.0, 0.3]	3	10
[0.3, 0.6]	5	8
[0.6, 0.9]	4	10
[0.9, 1.2]	4	5
[1.2, 1.5]	4	2
[1.5, 1.8]	3	2
[1.8, 2.1]	2	1
[2.1, 2.4]	4	3
[2.4, 2.7]	6	0
[2.7, 3.0]	1	0
……	……	……

由表 2 中两个数据库在不同区间内的语音数目分布可以发现，当 dist² 分布于[0, 1.8]内时，两类数据库的语音数目基本上保持一致；但当大于 1.8 时，语音分布数目发生起伏，甚至在柏林语音库中，没有相应的语音分布。因此，本文的聚类距离阈值取 1.35。

3.4　实验结果与分析

将聚类 PSO-LSSVM 回归模型与 LSSVM 回归模型、PSO-LSSVM 回归模型的预测结果相比较，根据预测值与标注值的数据分布以及实验性能指标来评价对 P、A、D 的预测能力。

3.4.1 数据分布

为更直观地比较三类回归模型的预测结果与标注值的数据分布，以对 TYUT2.0 数据库的预测结果为例，图 4~6 所示分别为回归模型对 P、A、D 维度的预测值与标注值的对比。

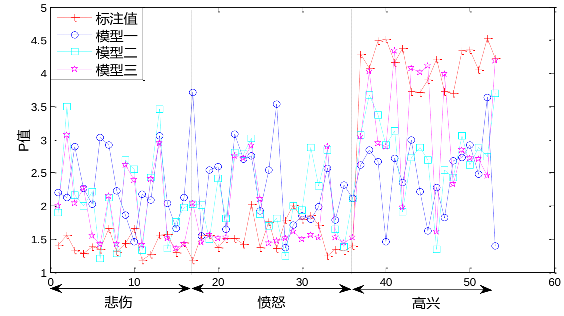


图 4　TYUT2.0 维度 P 的预测值与标注值对比

Fig. 4　Comparison of predictive values and tagged values of P in TYUT2.0 database

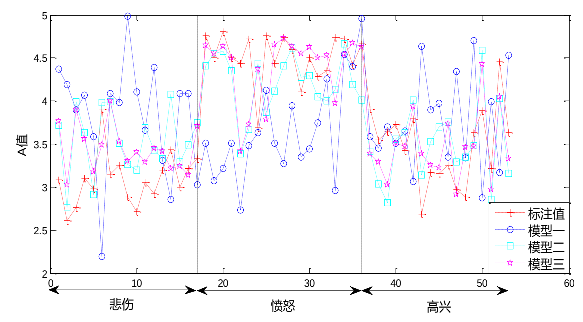


图 5　TYUT2.0 维度 A 的预测值与标注值对比

Fig. 5　Comparison of predictive values and tagged values of in TYUT2.0 database

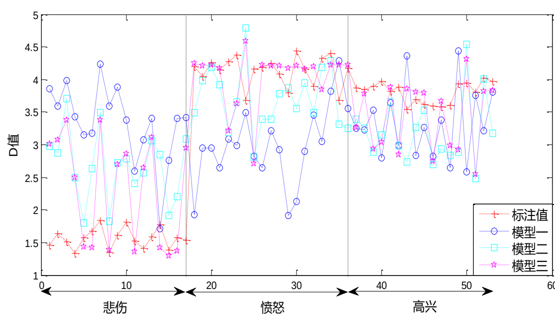


图 6　TYUT2.0 维度 D 的预测值与标注值对比

Fig. 6　Comparison of predictive values and tagged values of D in TYUT2.0 database

从图 4 可以看出，在悲伤和愤怒的语音中，模型二和模型三的预测值较相近，且相比模型一更贴近于标注的 P 值，而在高兴情感语音中，模型三的预测结果明显比其他两组模型更近于标注的 P 值分布；从图 5 可以看出，模型二和模型三预测结果的变化趋势相比模型一更贴近于标注 A 值的变化趋势，且模型三比模型二的预测值更接近于标注值、相对误差更小，可知相比于其他两组回归模型，模型三对每句语音的 A 值预测结果更为准确、稳定；从图 6 可以看出，模型三对三种情感情感维度 D 的预测结果更贴近于标注值，尤其在愤怒和高兴情感中，模型三预测结果离散度更接近于实际 D 值的离散度。说明模型三即聚类 PSO-LSSVM 回归模型的预测结果更接近于 P、A、D 标注值。

3.4.2 性能指标对比

实验性能指标使用平均绝对误差(MAE)及模型决定系数(R^2)作为模型的评价指标，再引入 Pearson 相关系数(r)评价模型预测值与标注值的变化趋势。MAE 越小越好，R^2 和 r 越接近 1 越好。其表达式分别为

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(\hat{y}_i - y_i)| \tag{7}$$

$$R^2=\frac{\sum_{i=1}^{n}(\hat{y}_i-u)^2}{\sum_{i=1}^{n}(y_i-u)^2},u=\frac{1}{n}\sum_{i=1}^{n}y_i \tag{8}$$

$$r=\frac{\sum_{i=i}^{n}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})}} \tag{9}$$

其中：n 为样本数；y_i 为实际标注值；$\hat{y}_i$ 为模型预测值。表 3、4 所示分别为 TYUT2.0 数据库和柏林语音库中三类回归模型对 P、A、D 预测的实验性能对比。

表 3　TYUT2.0 数据库的实验性能指标对比

Table 3　Comparison of experimental performance indexes in TYUT2.0 database

	P			A			D		
性能指标	模型一	模型二	模型三	模型一	模型二	模型三	模型一	模型二	模型三
Person 相关系数	0.44	0.51	0.63	0.68	0.69	0.75	0.59	0.62	0.74
模型决定系数	0.20	0.26	0.40	0.44	0.47	0.55	0.35	0.39	0.53
MAE	0.95	0.89	0.71	0.45	0.42	0.36	0.80	0.76	0.58

表 4　柏林语音库的实验性能指标对比

Table 4　Comparison of experimental performance indexes in EMO-DB

	P			A			D		
性能指标	模型一	模型二	模型三	模型一	模型二	模型三	模型一	模型二	模型三
Person 相关系数	0.45	0.48	0.63	0.67	0.67	0.74	0.96	0.96	0.98
模型决定系数	0.16	0.21	0.39	0.33	0.42	0.46	0.91	0.92	0.96
MAE	0.93	0.98	0.66	0.41	0.40	0.36	0.29	0.27	0.17

结合表 3、4，可以得出：

a)针对模型二与模型一对 P、A、D 维度的预测结果，从 Person 相关系数可以看出，在 TYUT2.0 数据库中模型二的相关系数相比模型一有一定程度的提高；从模型决定系数可以看出，在两类数据库中模型二的决定系数比模型一均有提高，说明模型二相比模型一对实验数据的拟合效果好；从 MAE 可以看出，在 TYUT2.0 数据库中，模型二的实验误差均有所减小，在柏林数据库中，对 A、D 的预测误差减小，但 P 的预测误差却增大。由此可知，模型二在对 A、D 维度的预测效果更稳定，说明 PSO 算法对 LSSVM 回归参数优化作用使预测效果有一定程度的改善。

b)针对模型三与模型二对 P、A、D 维度的预测结果，从 Person 相关系数可以看出，模型三对两类数据库的预测相关系数均比模型二有一定提高，说明模型三的预测结果变化趋势更加相近于标注值的变化趋势；从模型决定系数可以看出，模型三比模型二对数据的拟合作用更好；从 MAE 可以看出，在两类数据库中模型三的实验误差均比比模型二的误差更小。由此可知，模型三较模型二对 PAD 维度的预测性能更优，验证了情感聚类分析可以通过降低特征间的相关性而提高预测精度。

c)针对三类模型对两类数据库的 P、A、D 维度预测结果，可以看出，在 TYUT2.0 数据库中，三类回归模型对 A 维度的预测效果优于 P、D 维度；在柏林语音库中，三类回归模型对 D 维度的预测效果优于 P、A 维度；这与数据库的建立方式不同有关，柏林语音库是基于表演而录制的语音库，而 TYUT2.0 是通过截取广播剧的方式获得。虽然两类数据库在情感表达方面各有侧重，使得预测效果有所差异，但模型三对 P、A、D 维度的预测始终有不同程度的改进，而且与文献[15]对 P、A 预测结果相比较，模型三在两类数据库中对 P、A 预测的决定系数分别高于文献中利用 KNN 对 P、A 维度预测模型的决定系数 0.24 和 0.35。由此说明模型三的预测更优。

综上所述，模型三对情感维度 P、A、D 的预测性能不仅有较大幅度的提高，而且还适用于不同类型的情感数据库，说明聚类 PSO-LSSVM 模型对 P、A、D 的预测能力较强，且对数据的普适性好。分析其原因，PSO 算法对 LSSVM 模型回归参数的优化避免了参数主观选择的盲目性，使得预测结果有一定程度的改善；情感聚类分析通过将情感进行初步聚类，降低了不同情感特征之间的关联性对预测精度的影响，使其对 P、A、D 维度的预测性能得到改善；因此，聚类 PSO-LSSVM 回归模型的提出，不仅可以使回归过程更加客观，而且可以通过降低特征间的相关性，使该模型对 P、A、D 维度的预测更为准确。

4　结束语

为对情感从维度进行情感客观量化，本文提出了一种将 PSO 优化 LSSVM 与聚类分析相结合的预测方法，即聚类 PSO-LSSVM 回归模型，将其用于情感维度 P、A、D 的预测，并与 LSSVM、PSO-LSSVM 两种回归模型的预测结果进行对比。实验结果表明，聚类 PSO-LSSVM 回归模型对情感维度 P、A、D 的预测效果更好，这是因为该回归模型融合了 PSO 算法、情感聚类分析的优点，即 PSO 算法对回归参数的优化有效避免了参数主观选择的盲目性；情感聚类分析对 PSO-LSSVM 初次预测结果进行处理降低了情感特征之间的关联性，进而实现了更准确、智能的 P、A、D 预测。该结果能够更精确地展现出语音的情感维度，下一步可以通过预测语音样本在 PAD 空间中的坐标点，根据情感语音在 PAD 三维空间的分布分析其与基本情感在维度空间中的关系，深入解析任一语音情感状态的构成元素以及组成比例，为分析复杂情感的组成提供量化方法。

参考文献：

[1] 卜湛，伍之昂，曹杰，等. 在线评论情感计算与博弈预测 [J]. 电子学报, 2015, 43 (12): 2530-2535. (Bu Zhan, Wu Zhiang, Cao Jie, *et al.* Affective computing and game theory based prediction for online reviews [J]. Acta Electronica Sinica, 2015, 43 (12): 2530-2535.)

[2] 李春祥，丁晓达，叶继红. 基于混合蚁群和粒子群优化 LSSVM 的脉动风速预测 [J]. 振动与冲击, 2016, 35 (21): 131-136. (Li Chunxiang, Ding Xiaoda, Ye Jihong. Fluctuating wind velocity forecasting based on LSSVM with hybrid ACO & PSO [J]. Journal of Vibration and Shock, 2016, 35 (21): 131-136.)

[3] 耿立艳. 基于 GRA 与 KPCA 的 LSSVM 物流需求预测 [J]. 交通运输系统工程与信息, 2015, 15 (1): 137-142. (Geng Liyan. Forecast of logistics demand using LSSVM combining GRA with KPCA [J]. Journal of Transportation Systems Engineering and Information Technology, 2015, 15 (1): 137−142.)

[4] Zhao P C, Liu B, Gao W, *et al.* Multiple kernel least square support vector machine model for prediction of cement clinker lime content [J]. CIESC Journal, 2016, 67 (6): 2480−2487.

[5] Zhao Y, Niu R, Ling P, *et al.* Prediction of landslide deformation based on rough sets and particle swarm optimization-support vector machine

[J]. Journal of Central South University, 2015, 46 (6): 2324−2332.

[6] Yu H, Chen Y, Hassan S G, *et al.* Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO [J]. Computers & Electronics in Agriculture, 2016, 122: 94−102

[7] Gorjaei R G, Songolzadeh R, Torkaman M, *et al.* A novel PSO-LSSVM model for predicting liquid rate of two phase flow through wellhead chokes [J]. Journal of Natural Gas Science & Engineering, 2015, 24: 228-237.

[8] 宋静，张雪英，孙颖，等．基于 PAD 情绪模型的情感语音识别 [J]. 微电子学与计算机, 2016, 33 (9): 128-131. (Song Jing, Zhang Xueying, Sun Ying, *et al.* Emotional speech recognition based on PAD emotion model [J]. Microelectronics & Computer, 2016, 33 (9): 128-131.)

[9] 王丽．V-A 空间连续维度情感预测方法研究 [D]. 镇江：江苏大学，2015. (Wang Li.Research on emotional prediction method of V-A spatial continuous dimension [D].Zhenjiang: Jiangsu University,2015.)

[10] 褚玉伟，罗晓博，屈珂，等. DBSCAN 和 K-Means 混合聚类的牙齿特征自动识别 [J]. 计算机辅助设计与图形学学报，2018 (7): 1276-1283. (Chu Yuwei, Luo Xiaobo, Qu Ke, *et al.* DBSCAN and K-means hybrid clustering based automatic dental feature detection [J]. Journal of Computer-Aided Design & Computer Graphics, 2018 (7): 1276-1283.)

[11] 谭飞刚，刘伟铭，黄玲，等．基于加权欧氏距离度量的目标再识别算法 [J]. 华南理工大学学报 :自然科学版 ，2015 (9): 88-94. (Tan Feigang, Liu Weiming, Huang Ling, *et al.* Object re-identification algorithm based on weighted euclidean distance metric [J]. Journal of South China University of Technology :Natural Science Edition , 2015 (9): 88-94.)

[12] 张雪英，张婷，孙颖，等．情感语音数据库优化及 PAD 情感模型量化标注 [J]. 太原理工大学学报，2017, 48 (3): 469-474. (Zhang Xueying, Zhang Ting, Sun Ying, *et al.* Emotional speech database optimization and quantitative annotation based on PAD emotion model [J]. Journal of Taiyuan University of Technology, 2017, 48 (3): 469-474.)

[13] Burkhardt F, Paeschke A, Rolfes M, *et al.* A database of German emotional speech [C]// Proc of INTERSPEECH 2005-Eurospeech, European Conference on Speech Communication and Technology. Lisbon, Portugal:DBLP, 2005: 1517-1520.

[14] 姚慧，孙颖，张雪英．情感语音的非线性动力学特征 [J]. 西安电子科技大学学报，2016, 43 (5): 167-172. (Yao Hui, Sun Ying, Zhang Xueying. Research on nonlinear dynamics features of emotional speech [J]. Journal of Xidian University, 2016, 43 (5): 167−172)

[15] Giannakopoulos T, Pikrakis A, Theodoridis S. A dimensional approach to emotion recognition of speech from movies [C]// Proc of IEEE International Conference on Acoustics, Speech and Signal Processing. 2009: 65-68.